

SYSTEMS, METHODS AND COMPUTER PROGRAM PRODUCTS FOR  
ASSOCIATING DYNAMICALLY GENERATED WEB PAGE CONTENT  
WITH WEB SITE VISITORS

Field of the Invention

The present invention relates generally to Web pages and, more particularly, to Web pages containing dynamically generated content.

5

Background of the Invention

The Internet is a worldwide decentralized network of computers having the ability to communicate with each other. The Internet has gained broad recognition as a viable medium for communicating and interacting across multiple networks. The World Wide Web (Web) was created in the early 1990's and is comprised of server-hosting computers (Web servers) connected to the Internet that have hypertext documents or Web pages stored therewithin. Web pages are accessible by client programs (i.e., Web browsers) utilizing the Hypertext Transfer Protocol (HTTP) via a Transmission Control Protocol/Internet Protocol (TCP/IP) connection between a client-hosting device and a server-hosting device. While HTTP and hypertext documents are the prevalent forms for the Web, the Web itself refers to a wide range of protocols including Secure Hypertext Transfer Protocol (HTTPS), File Transfer Protocol (FTP), Gopher, and content formats including plain text, Extensible Markup Language

09207945-120998  
(XML), as well as image formats such as Graphics Interchange  
Format (GIF) and Joint Photographic Experts Group (JPEG).

Specific to the Web, a Web server is a computer  
program (housed in a computer) that serves requested Web  
5 pages and files. A Web client is a requesting program  
associated with a user. A Web browser is an exemplary Web  
client for use in requesting Web pages and files from Web  
servers.

A Web site is conventionally a collection of Web  
10 pages and files related to a particular subject that  
includes a beginning file called a home page. A large Web  
site may reside on a number of geographically-dispersed Web  
servers. The Web site of the International Business Machines  
Corporation (www.ibm.com), for example, consists of  
15 thousands of Web pages and files spread out over various Web  
servers in locations world-wide.

As is known to those skilled in this art, a Web  
page is conventionally formatted via a standard page  
description language such as HyperText Markup Language  
20 (HTML), which typically contains text and graphics  
references and can reference sound, animation (e.g., via  
programs and/or scripts), and video data. HTML provides for  
basic document formatting and allows a Web content provider  
to specify hypertext links (typically manifested as  
25 highlighted text) to other Web servers and files. When a  
user selects a particular hypertext link, a Web browser  
reads and interprets an address, called a Uniform Resource  
Locator (URL) associated with the link, connects the Web  
browser with the Web server at that address, and makes an  
30 HTTP request for the file identified in the link. The Web  
server then sends the requested file to the client in HTML  
format which the Web browser interprets and displays to the  
user.

Exemplary Web browsers include Netscape Navigator® (Netscape Communications Corporation, Mountain View, CA) and Internet Explorer® (Microsoft Corporation, Redmond, WA). Web browsers typically provide a graphical user interface for retrieving and viewing information, applications and other resources hosted by Web servers.

In the business world of yesteryear, shopkeepers often knew their customers by name. Typically, a shopkeeper knew his/her customers' preferences with respect to the products and services offered by the shopkeeper. By knowing information about his/her customers' preferences, a shopkeeper could more adequately meet customer needs and ensure greater business success. For the same reason today, Web site owners would like to learn about the preferences of their customers. Unfortunately, Web site owners typically do not personally know customers who access their Web sites via the Internet.

Many Web pages contain dynamically generated content. Dynamically generated content allows Web site owners to select the content for a delivered Web page, as well as manipulate the placement of that content within the Web page, for various purposes. It would be desirable to collect information about the effectiveness of content selection and placement each time a Web visitor comes to a Web site, for example by measuring the length of time a visitor views particular content within a Web page. By knowing this type of information, a Web site owner can determine the efficacy of content presentation on the Web site, and can tailor content to particular users. Unfortunately, current Web servers may not have the ability to collect information about the content that a Web site visitor views.

### Summary of the Invention

In view of the above discussion, it is an object of the present invention to facilitate personalized content delivery to Web site visitors.

5 It is another object of the present invention to enhance the efficacy of information delivery to Web site visitors.

10 It is another object of the present invention to facilitate collecting information about the preferences of Web site visitors.

It is another object of the present invention to facilitate collecting information about dynamically generated Web content requested by Web site visitors.

15 These and other objects of the present invention are provided by systems, methods and computer program products for associating dynamically generated Web page content with a user who requests a Web page from a Web server. Upon receiving a user request for a Web page via a Web client in communication with the Web server, a Web  
20 server records the user request within a Web server log. The requested Web page is then dynamically generated using a template for the requested Web page and one or more content objects. Exemplary content objects include, but are not limited to, text files, audio files, video files, image  
25 files, and hyperlinks. Each content object has a respective unique identifier associated therewith. The stored record of the user request is then appended with each identifier for respective content objects included within the displayed Web page.

30 The present invention is advantageous because a plurality of stored user request records appended with content object identifiers, according to the present invention, can be analyzed to determine the type of content a user views as well as the time spent viewing the content.

866027" 54620260

Furthermore, the efficacy of certain content within a viewed Web page can be determined via the present invention. For example, if particular content is generally ignored by users (e.g., users do not follow hyperlinks contained within the content), a Web site owner can relocate or reformat the particular content.

The present invention stems from the realization that requests for Web pages made by a Web site visitor are conventionally accompanied by some type of user information, such as who the visitor is, the visitor's location, the type of Web browser utilized, a time stamp, and the like. This user information is conventionally stored within a Web server log file. Unfortunately, user information provided with a Web page request may not provide a Web site owner with relevant information about a visitor's preferences. Unless visitors are requested to fill out a form detailing their preferences, Web site owners typically do not have a way of gathering such information about their Web site visitors.

#### Brief Description of the Drawings

**Fig. 1** schematically illustrates operations according to the present invention for associating dynamically generated Web page content with users requesting Web pages from a Web server.

**Fig. 2** schematically illustrates a system in which the present invention can be implemented.

#### Detailed Description of the Invention

The present invention now is described more fully hereinafter with reference to the accompanying drawings, in which preferred embodiments of the invention are shown. This invention may, however, be embodied in many different forms and should not be construed as limited to the embodiments

set forth herein; rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. Like numbers refer to like elements throughout.

5           As will be appreciated by one of skill in the art, the present invention may be embodied as a method, data processing system, or computer program product. Accordingly, the present invention may take the form of an entirely hardware embodiment, an entirely software embodiment, or an  
10   embodiment combining software and hardware aspects. Furthermore, the present invention may take the form of a computer program product on a computer-usable storage medium having computer-usable program code means embodied in the medium. Any suitable computer readable medium may be  
15   utilized including hard disks, CD-ROMs, optical storage devices, or magnetic storage devices.

Computer program code for carrying out operations of the present invention is preferably written in an object oriented programming language such as JAVA®, Smalltalk or  
20   C++. However, the computer program code for carrying out operations of the present invention may also be written in conventional procedural programming languages, such as the "C" programming language, or in a functional (or fourth generation) programming language such as Lisp, SML, or  
25   Forth. The program code may execute entirely on a web server, or it may execute partly on a web server and partly on a remote computer (i.e., a user's web client). In the latter scenario, the remote computer may be connected to the web server through a LAN or a WAN, or the connection may be  
30   made to an external computer (for example, through the Internet using an Internet Service Provider).

The present invention is described below with reference to flowchart illustrations of methods, apparatus (systems), and computer program products according to an

embodiment of the invention. It will be understood that each block of the flowchart illustrations, and combinations of blocks in the flowchart illustrations, can be implemented by computer program instructions. These computer program instructions may be provided to a processor of a general purpose computer, special purpose computer, or other programmable data processing apparatus to produce a machine, such that the instructions, which execute via the processor of the computer or other programmable data processing apparatus, create means for implementing the functions specified in the flowchart block or blocks.

These computer program instructions may also be stored in a computer-readable memory that can direct a computer or other programmable data processing apparatus to function in a particular manner, such that the instructions stored in the computer-readable memory produce an article of manufacture including instruction means which implement the function specified in the flowchart block or blocks.

The computer program instructions may also be loaded onto a computer or other programmable data processing apparatus to cause a series of operational steps to be performed on the computer or other programmable apparatus to produce a computer implemented process such that the instructions which execute on the computer or other programmable apparatus provide steps for implementing the functions specified in the flowchart block or blocks.

Referring now to **Fig. 1**, operations according to the present invention for associating dynamically generated Web page content with a user who requests a Web page from a Web server are illustrated. Initially, a Web server receives a user request for a Web page via a Web client in communication with the Web server (Block 100). A record of the user's Web page request is stored within a Web server log (Block 200). This log record includes the URL of the

09207945 120998

requested Web page, the requesting user's identity, and the time stamp of the request. The Web server then dynamically generates the requested Web page using a template and one or more content objects (Block 300). Each content object used to generate the requested Web page has a unique identifier associated therewith. The generated Web page is then served to the Web client (Block 400). The unique identifier for each content object included within the served Web page is appended to the stored user request record (Block 500). Each time a user requests additional content, the respective identifiers associated with the requested content object(s) are appended to the respective user request record, along with the time stamp of that content request. Finally, because time stamps are conventionally a part of a user request record, the time a user spends viewing particular content also can be determined. When a user requests a new Web page, the server records the time of this request with the associated request record for the user's previous Web page request. In this way, a user is assumed to have viewed content from the time it was delivered (and a corresponding entry is made to the request record) to the time that the user initiates a new Web page request (and the request time is recorded in that request record).

Stored records of user requests can be analyzed to determine user preferences with respect to Web page content (Block 600). The choices of Web pages requested by the user, combined with information about what content was presented on each requested Web page and the amount of time spent viewing that content, can provide insight into the preferences of a user. Furthermore, the efficacy of certain content within a viewed Web page can be determined via the present invention. For example, if particular content is generally ignored by users (e.g., users do not follow hyperlinks and, consequently, request corresponding Web



pages, contained within the content), a Web site owner can relocate or reformat the particular content.

Referring now to Fig. 2, a system within which the present invention may be implemented is schematically illustrated. A Web client 10 is in communication with a Web server 12 via a computer network, such as the Internet 14. The Web server 12 includes a logging facility 16 for storing records of user requests for Web pages. Web server logging facilities are well understood by those skilled in this art and need not be described further herein.

The Web server 12 is configured to dynamically generate a requested Web page 30 using a template 18 and one or more Web page content objects. The template 18 defines the layout of content objects within the Web page 30 when displayed within a Web client (i.e., Web browser). Each element of a Web page, including, but not limited to, divisions, sections, headings, paragraphs, images, lists, tables, and hyperlinks, may be represented by a content object; in addition, a content object may include audio and video files. It is understood, however, that a single content object may represent one or more of these Web page elements. The illustrated template 18 includes a plurality of fields 17a-17e that indicate where a respective content object is to be displayed within a generated Web page.

In the illustrated embodiment, content objects are stored within a database 19 that is accessible by the Web server 12. However, it is to be understood that content objects 20 can be stored in a file system, as well. When a user request is received for the Web page 30, the Web server 12 retrieves the template 18 and content objects and generates the Web page 30. Dynamic generation of Web pages using templates and content objects is known to those skilled in this art and need not be described further herein.

Each content object 20 in the generated Web page 30 has a unique identifier (1-5) associated therewith, as illustrated. Content objects are preferably assigned a unique identifier prior to receiving a request for a Web page. This facilitates saving processor time and speed during run time when Web pages are dynamically generated. Preferably a unique identifier is generated using a hashing function. As is known by those skilled in this art, hashing is the transformation of a string of characters into a usually shorter fixed-length value or key that represents the original string. A well known hashing function that is particularly suitable for generating unique identifiers for content objects is the Message Digest 5 (MD5) hashing function. MD5 was authored by Ronald L. Rivest, Massachusetts Institute of Technology, Laboratory for Computer Science, Cambridge, MA, and is described in the Network Working Group Request for Comments 1321 (RFC 1321), Information Sciences Institute, University of Southern California, Marina del Rey, CA, April, 1992, which is incorporated herein by reference in its entirety. In alternative embodiments of the present invention, the content object's unique identifier may be formed as a URL or Uniform Resource Identifier (URI), optionally appended with additional unique identifying information.

The foregoing is illustrative of the present invention and is not to be construed as limiting thereof. Although a few exemplary embodiments of this invention have been described, those skilled in the art will readily appreciate that many modifications are possible in the exemplary embodiments without materially departing from the novel teachings and advantages of this invention. Accordingly, all such modifications are intended to be included within the scope of this invention as defined in the claims. In the claims, means-plus-function clause are

intended to cover the structures described herein as  
performing the recited function and not only structural  
equivalents but also equivalent structures. Therefore, it is  
to be understood that the foregoing is illustrative of the  
5 present invention and is not to be construed as limited to  
the specific embodiments disclosed, and that modifications  
to the disclosed embodiments, as well as other embodiments,  
are intended to be included within the scope of the appended  
claims. The invention is defined by the following claims,  
10 with equivalents of the claims to be included therein.

0920794-12099  
SECRET - 54640260